A Study on Automated Receptionists in a Real-World Scenario

Wolter, Ralf; Hindriks, Koen V.; Samur, Dalya; Jonker, Catholijn M.

published in
Advances in Practical Applications of Agents, Multi-Agent Systems, and Trustworthiness. The PAAMS Collection
2020

DOI (link to publisher)
10.1007/978-3-030-49778-1_27

document version
Publisher's PDF, also known as Version of record

citation for published version (APA)
Wolter, R., Hindriks, K. V., Samur, D., & Jonker, C. M. (2020). A Study on Automated Receptionists in a Real-World Scenario. In Y. Demazeau, T. Holvoet, J. M. Corchado, & S. Costantini (Eds.), *Advances in Practical Applications of Agents, Multi-Agent Systems, and Trustworthiness. The PAAMS Collection: 18th International Conference, PAAMS 2020, L'Aquila, Italy, October 7–9, 2020, Proceedings* (pp. 340-352). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 12092 LNAI). Springer. https://doi.org/10.1007/978-3-030-49778-1_27

A Study on Automated Receptionists in a Real-World Scenario

Ralf Wolter[2,3](✉), Koen V. Hindriks[1], Dalya Samur[3], and Catholijn M. Jonker[2]

[1] Vrije Universiteit, Amsterdam, The Netherlands
k.v.hindriks@vu.nl
[2] Delft University of Technology, Delft, The Netherlands
{r.c.Wolter,c.m.jonker}@tudelft.nl
[3] ING, Amsterdam, The Netherlands
ralf.wolter@ing.com, dalyasamur@gmail.com

Abstract. The commercial availability of robots and voice-operated smart devices such as Alexa or Google Home have some companies wondering whether they can replace some current human interactions by using these devices. One such area of interaction is at the reception desk. While both platforms can offer the necessary interaction features to take on the task of an automated receptionist, the question remains as to which platform actual visitors would prefer - body or no body? To this end we created a receptionist agent that can receive visitors with an appointment, presented as either an embodied robot or a disembodied smart display. The agent uses common commercial products and services, and was tested in a real-world environment with real visitors.

The results show no significant difference in visitor preference for either platform.

Keywords: Human-robot interaction · Embodiment · Social agents · Dialogue management · Automated receptionist

1 Introduction

The user experience of people visiting bank branches has not changed much since the inception of the savings bank. A customer enters, waits in line, and states their needs to an employee. If it is a routine request the employee can help directly and the visit ends there. For more important meetings the customer can make a private appointment with a more specialised employee.

While over the years there has been little change in the bank visiting experience, ever-improving communication technology steadily reduced the need for customers to visit the bank branch in person [18]. From telegraph, to telephone, and now internet, increasingly secure and reliable channels allow many routine actions to be performed without visiting the bank. Recent years have even given rise to banks without any branches at all.

As more and more of the routine tasks are taken over by internet and mobile banking applications, the few visits left become more and more important to

Y. Demazeau et al. (Eds.): PAAMS 2020, LNAI 12092, pp. 340–352, 2020.
https://doi.org/10.1007/978-3-030-49778-1_27

both bank and client [21]. The decisions made during the visit could have a large impact on the clients personal future. From the perspective of the bank, it is one of the few times that it has direct face-to-face contact with the customer. As such it is one of the few opportunities to deepen the relationship.

To make the most of these scarce moments of contact, some companies are experimenting with new types of offices [8,9]. Some have replaced counters and stuffy meeting rooms with sitting areas decorated to feel like small living rooms. This is done in order to make a visitor feel welcome and at home. Similarly the receptionist plays a huge role in the customer experience. Making the customer be at ease and engaged and delivering a personalized experience is essential for a good customer relationship [4].

Even when entering a building with such a homely atmosphere, customers could still feel uncertain and uncomfortable if they are unfamiliar with the environment. It is then paramount to engage with these visitors as soon as possible to minimize those feelings [13]. Typically this is the role the receptionist fulfills.

With the recent advancements in robotics and AI, companies have been experimenting with automated assistants and receptionists to further improve customers experience [10,23]. The popularity of humanoid robots such as Softbank's Pepper [20] and smart displays such as Amazon Echo Show [1] make these obvious candidates for automating a receptionist. Speech recognition and generation capabilities allow for direct spoken dialogue between receptionist and visitor, while the electronic nature of the receptionist allows for sending and retrieval of relevant information in the background. With the wide variety of capabilities of these technologies, it is important to investigate which of their features contribute to an effective user experience when they are used as receptionists. One of the most striking difference between the two above-mentioned technologies is that a robot, such as Pepper has a humanoid robot body, while others, such as Alexa, use only voice for the interaction. Given these differences, especially with regard to the impact on deployment cost, it is important to investigate how important it is to give an automated receptionist a body. The availability of these tools raises the question of how important the embodiment is when replacing a counter with an automated receptionist. That is, if, in a real-life setting, we were to replace a reception desk with an automated agent, how important is it to give the agent a body? More precisely we studied:

Research Question: To what extent does providing a body for an automated receptionist contribute to a better user experience?

To this end we created a prototype agent that could handle receiving guests who have appointments. It engages a visitor in a verbal dialogue based on observed dialogues between receptionists and visitors. This agent was given two different housings, a Pepper robot, and a smart display.

The prototype was tested in a real life setting at a recruitment office of ING, a large international banking organization with its head office in the Netherlands. Visitors were almost exclusively job applicants coming in for an interview. Upon entry they were confronted with either a robot or a smart display. Their experiences were analysed in order to determine the effect of a robot body on user experience.

2 Related Work

The value of physicality in interaction fascinated humans even before the invention of robots. Interactions with robots or avatars have been studied and compared to other forms of communication. In these comparative studies some physical aspect of the robot is compared to virtual (visual) alternatives (for examples, see [3,7,14,15,22,24]). In most cases the virtual agent is a digitized avatar of the robot, so in essence an animation of the same robot on a screen.

The social settings studied vary from entertainment and games to informational and directive interactions. These are more extended tasks to the to the task in this research. Some compare an information display to a robot where a more traditional interface is put against a fully physical social robot. Hence, the mode of interaction also differs between the two options. In contrast, the present research takes an between path in which the interaction is performed through spoken dialogue by all parties, but there is no graphic representation of the assistant just a disembodied voice.

These comparisons are also generally studied in artificial lab scenarios. An exception is Lee et al. [16], who studied interactions with a receptionist robot by simple placing it in a busy area. The studies all reported a positive benefit to the physical robot above the virtual in their dependent variable. Bainbridge et al. and Lee et al. [3,15] listed more positive user ratings as a benefit.

Robots have been studied in real-world contexts, e.g., [6,17]. However, these studies were not comparative in nature, but studied the feasibility of having robots act effectively in the chosen situations. These studies have involved not only highly specialised robots such as SAYA [12] and Telenoid [19], but also commercial robots such as Nao, Pepper, Sophia [11] and Emiew3 [2].

To our knowledge, the present experiment is the first in which two automatized receptionists, with body and no body options, welcomed participants in the same experimental context which enables comparison of user experience with a commercial robot and a commercial audio platform.

3 Research Design

To answer the research questions stated in Sect. 1, the experimental design focuses on evaluation of the agent-user dialogue and user experience. Dialogue evaluation relies on transcription logs and classification of the utterances by human experts. User experience is measured by social robotics questionnaires completed by the participants after the interaction.

3.1 Groups and Participants

The participants were all job applicants for positions ranging from technical to human resources oriented. They were asked whether they wanted to participate in the experiment at the invitation to the interview. Due to the sensitive nature of a job interview, participants were informed that participation was voluntary and optional, and had no bearing on the interview.

Forty-six applicants agreed to participate in the experiment. 12 of the participants were female and 34 were male. They were between 22 and 36 years of age with an average age of 27.2. Five of the participants listed high school as their highest completed education, two had received their PhD and thirty-nine were university graduates.

For the experiment participants were divided into two groups. One group interacted with the agent housed in a Pepper robot, the other with the agent housed in the smart display setup. During interview days, the receptionist was setup in the morning and all visitors for the day were received with the same setup. All participants on the same day therefore encountered the same setup and were placed in one of the two groups based on this. Data were not collected from non-participants.

3.2 The Location

The experiment took place in an old farm remodeled as an office. Aside from the bank's human resource department, the building also houses the corporate academy used for training recruits.

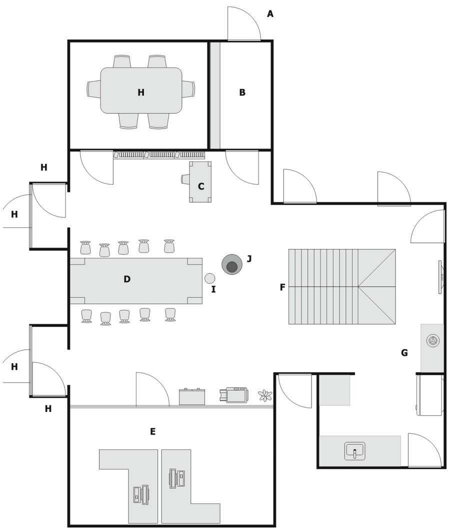

Fig. 1. Floor plan of the reception area.

All visitors enter the building through the main entrance at A and B, shown in Fig. 1. A receptionist located at C receives guests here. Guests are directed to a nearby area D to await an employee.

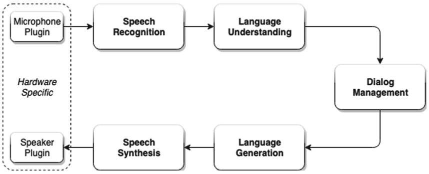

Fig. 2. Agent architecture

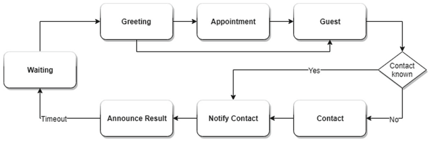

Fig. 3. Reception dialogue states

For the experiment, the desk at C where the receptionist would normally be sitting was removed. The human receptionist took seat nearby in area E to answer the phone, but not interact with guests. One exception was that when the agent had a fatal breakdown in the dialogue, it asked the visitor to wait and notified the receptionist instead of a proper appointment contact with the appropriate employee.

The agent housing was placed in front of the entrance at location J. A sign was hung on a support beam next to it asking visitors to report to the robot or smart display. The sign gave no instructions as to how this should be accomplished.

3.3 The Agent

The goals for the agent were twofold: First to be able to run different task-based dialogues as needed and secondly to execute the same dialogue independent of the hardware used. The agent is designed as a generic lightweight simple dialogue system. Dialogues can be specified as a simple state machine in a simple description format. This method was preferred above more advanced techniques for its clear predictability and control.

The architecture follows the classical dialog system approach as shown in Fig. 2. The pipeline is implemented as an asynchronous observable stream, each component reacting to incoming messages and producing the next message in the stream. The microphone and speaker plugins provide hardware specific hot spots where the behaviour can be customized to the housing of the agent. Speech recognition, language understanding and speech synthesis are wrapper components linking to commercial cloud services for the related tasks. Dialogue management implements a finite state machine-based execution environment for specifications of states and events.

It is common for spoken dialogue hardware to have some form of speech recognition and speech synthesis built-in. The quality of these pre-packaged solutions varies. So as not to let the quirks of each system influence the results, the designed agent skips the internal speech recognition, language understanding and speech synthesis. Instead, the agent components make use of the Google services for speech, dialogue flow and text-to-speech.

The dialogue manager handles incoming events, outgoing actions and session storage. A specification creates mappings between events and one or more actions based on the current state. Events include activation, semantic input as the result of the language understanding component, errors and timer notifications. Actions can be semantic output, state transition, store input, clear store, and execute plugin. One composite action is defined for selectively executing another action based on an expression. Plugins contain domain specific code, mainly used to communicate with external systems. This set of primitives creates a small specification language sufficient to handle simple task-based dialogues.

3.4 Agent Housing

For this experiment two housings were used for the agent. One is a Softbank Pepper robot and the other custom hardware representing the various smart displays on the market such as Amazon Echo and Google Home. Both have some lightweight plugins written for the specific requirements of the hardware.

In order to ensure that both setups would behave much the same and to control for aspects such as pitch and gender in the experiment, the speech generation and speech synthesis used in both setups are the same. Only the bodily behaviour differs between the setups.

Pepper. Pepper is a humanoid robot created by Softbank. It normally speaks with a high pitched child-like voice and makes contextual gestures. To keep the contextual gestures but speak in the agents (more mature) voice the plugin uses the normal speaking method with volume set to zero, while at the same time playing the audio stream generated by the agent. Audio from Pepper's speakers is directly streamed to the agent. In all other aspects, the default autonomous life feature of Pepper was used.

Smart Display. A custom setup was created for the smart display. Commercially available alternatives have closed-in aspects that allow direct streaming of audio and voice information. This would prevent the use of the agent's voice components. The custom setup used instead a Raspberry Pi Model B with audio capabilities. The plugins used are the generic Linux audio plugins for the agent.

3.5 Dialogue Specification

The model for the reception dialogue was based on a set of observations of human receptionists in action. The receptionists were asked to perform their task as normal while researchers observed. The conversations between receptionist and guest were also recorded.

The result was a set of recordings that give a general idea of how a human approaches this task. These recordings were transcribed and analysed to create a model for the dialogue as the agent should perform it.

In the diagram, shown in Fig. 3, boxes represent states derived from the transcripts. The sequence of events in the observed data always followed a fixed order. First the receptionist greets the guest, then guest are asked whether they have an appointment, followed by asking for the guest's name. Finally the receptionist asks for the contact name, after which the contact is called and the contact's answer is announced.

The data set showed three possible variations of this order. Some guests would tell the receptionist they had an appointment in response to the greeting, by passing the appointment query state. Other guests mentioned the contact's name during the greeting, allowing the contact query state to be skipped. The last variation combined the other two.

The "Greeting" state expects either a greeting intent or an appointment intent for the user. The "Appointment", "Guest" and "Contact" states query the user for the specific information and store this in the session. The "Notify Contact" state makes an external call and "Announce Result" state vocalizes the contact response. These states do not expect any user interaction. The decision state "Contact Known" is specified as a state handling only its activation events with guards to differentiate for the data in the session. The waiting state resets the session.

The model differs from the transcripts gathered from the human-human dialogues in one aspect. When notifying the contact the receptionist would pick up the phone and call the contact. The guest could see and hear this activity. For the dialogue system sending a message to the contact is silent and in the background, providing no clue to the guest as to what is going on. For this reason the system announces to the guest that it will notify the contact.

Repair Strategies. In one transcript shows the guest walked by the receptionist while giving a greeting in passing. Because the initial model was based on only a small set of dialogues, it is possible that a visitor could utter phrases the system has not yet heard. This indicates the need for recovery strategies.

A number of recovery strategies have been identified over the years. These range from repeating the question to giving full help information. Bohus et al. [5] showed that humans mostly employ asking the user to repeat and briefly going through the options of what the user could say.

The recovery strategy employed is a straightforward sequence of strategies. It first asks the users to repeat themselves (AREP), then makes a short suggestion of what to say (TYCS) and finally gives a more complete explanation of the user's options (YCS).

If all recovery actions (AREP, TYCS, YCS) fail to reach the intended result, the conversation is considered to have broken down and the system executes the fallback strategy. The fallback strategy is asking the guest to take a seat in the waiting area and notifying an employee to handle the task in person.

The normal strategy described above is used for states that require important information for the dialogue, such as appointment, guest and contact. A state such as greeting is considered less critical and has a more lenient recovery strategy in which the greeting is first repeated, if the system still does not understand after that it simply moves on to the next state.

3.6 Running the Experiment

The experiment ran for a total of 24 days. All participants were notified a week before about the data collection for the experiment and gave permission for their reception conversations to be recorded.

When a participant entered there was no other contact before engaging with the setup. Participants were expected to figure out where and how to interact with the setup on their own initiative. The only hint was a sign asking them to present themselves to the setup. Researchers observed the interaction of the participants with the setup. The agent automatically recorded and logged the conversation.

After the participant had taken a seat in the waiting area, a researcher approached the participant with the questionnaire. The employee contacts had been asked to wait ten minutes before approaching the participant.

3.7 The Measures

The conversation logs contained the textual transcript of the conversation as made by the agent, its' internal decisions, as well as the entire audio file of the conversation. The audio file was transcribed manually by a research assistant and afterwards compared with the logs for error analysis.

The questionnaires, offered to the participants, consist of three parts. A generic user experience part, namely "UX-questionnaire" [26], a part devoted to the specifics of embodiment, namely the "AttrakDiff questionnaire" [25], and a small number of open-ended questions.

The "UX-questionnaire" by Weiss and colleagues measures user experience focusing on five dimensions: Emotion (EM), Embodiment (EB), Feeling of Security (FoS), Human-oriented Perception (HOP), and Co-experience (COE). Each factor is assessed with five statements which were rated on a 7-point Likert scale.

AttrakDiff is a validated standardized questionnaire to get general insight on UX. This questionnaire is designed to measure four aspects: pragmatic quality of the system (PQ), hedonic quality identification (HQ-I), hedonic quality stimulation (HQ-S), and attractiveness (ATT). PQ describes the usability of the system. HQ-I is the extent to witch the user can identify with the system, and HQ-S is the extent to witch the system can stimulate functions and interactions. ATT describes the perceived quality of the system. This questionnaire consists of word-pairs, e.g. disagreeable - likable. Participants rate on a 7-point Likert scale, where negative word anchor is -3, and the positive anchor pole is 3.

Each of the two setups had its own version of the questionnaire, with the same questions but with the descriptions changed to correctly describe the setup. For the Pepper setup, the questions referred to "the robot", while for the display setup they referred to "the smart display". In all other aspects the two versions were exactly the same.

4 Results

After the initial testing 46 participants were received by the two setups. Of these 43 successfully completed the interaction. The participants included 12 females and 34 males and were between 22 and 36 years of age. In total 18 subjects (5 female, 13 male) interacted with the smart display setup, and 25 (5 female, 20 male) with the Pepper setup.

The error rate for all trials was 0.27. When normalized to the setups, the robot's rate was 0.32, while the smart display's rate was 0.19. The overall recovery rate was 0.94, 0.65 rate after one action, 0.87 after two actions, and 0.94 after the third action. During the course of the experiment, the fallback strategy was executed three times, a 0.06 breakdown rate.

The results for the AttrakDiff questionnaire are presented in Table 1. The values range from -3 to 3 in this scale; hence a mean value above 0 indicates a positive view of the experience, while below 0 indicates a negative view. Overall, the participants evaluated the experience as neutral and there were no significant differences between the two setups.

For the UX-Questionnaire, the values range from 1 to 7. The higher the value, the more positive the participants rated their experience. In this range 4.0 can be considered neutral with values lower than that negative and higher values positive. Table 2 shows a slight positive trend for "Feeling of Security" (FoS) while "Co-experience" (CoE) can be considered slightly negative. Differences between the robot setup and the smart display were not statistically significant.

5 Discussion

During the experiment 46 participants were received. These were serious applicants for job positions who were invited for interview as part of the normal operations of the human resources department. The procedure for the interviews was exactly the same as normal, except for their reception when entering the building, where the experiment was performed. This setup was dependent on the available pool of participants that were invited by the department.

Table 1. Results AttrakDiff

	ATT	HQI	HQS	PQ
Display Mean	0.532	0.357	0.873	0.794
Display S.D.	1.077	0.9754	1.030	0.8405
Robot Mean	0.362	0.259	1.167	0.490
Robot S.D.	0.8802	0.6564	0.7118	0.8619
t value	−0.069	−0.334	0.855	−0.983
p value	0.945	0.7404	0.398	0.332

Table 2. Results UX questionnaire

	COE	EB	EM	FoS	HoP
Robot Mean	2.589	4.390	4.470	5.005	4.227
Robot SD	0.7614	0.8872	0.6400	0.8126	0.8486
Display Mean	3.069	4.111	4.306	4.733	4.589
Display SD	1.143	1.058	0.8250	0.7388	0.6668
t value	−1.423	1.248	1.436	1.299	−1.076
p value	0.163	0.219	0.159	0.201	0.289

It is likely the participants were some form of stress during the experiment as they had their interview afterwards. This may have some influence on the results. We did not control for this variable. An topic requiring an appointment at a financial institution is almost by definition also a high state meeting as it likely involves a significant amount of financial resources. Therefore we made an assumption the participants would yield results similar to the targeted audience.

This brings us to the next question of how participants experienced the two platforms we tested. The questionnaire results show the users to be neither extremely positive, nor extremely negative about their experience.

In the present user experience results, we did not find any significant difference between the two setups. Other researchers have reported better user ratings for embodied agents [3,15]. We speculate the differences could be the result of the interaction length (about 45 seconds), task related, and/or the limited number of subjects. Looking at the magnitude of the effect further research with much more participants would be needed. A conservative estimate indicated about twenty times the current number is needed.

Such a followup would require a significant investment of time and money. And while we would consider a conclusive answer to the question whether embodiment actually influences the user experience in a financial reception setting to be valuable, other variables to be investigated could be considered at least as valuable.

Another aspect to consider is that the two setups are far from equal in regard to price. A single Pepper unit costs as much a hundred smart display units. This combined with the uncertainty of the actual benefit of embodiment in this context would make such an investment risky at best.

In addition, while the hardware may differ, the actual agent was the same for both setups. This means further development of AI behind the agent is actually independent of the physical representation. This gives us the opportunity to use the inexpensive smart displays to test and experiment. In time, robots could become much cheaper and reliable. Making the transition at that time should still be possible as long as the hardware/software duality remains observed.

One intriguing possibility of the separation between hardware and software is the effect of mobility on the user experience. As the software agent can switch between robot and display for the experiment, it is also conceivable to switch between hardware on the fly during conversation. This would allow the agent to be divorced from its fixed location in the lobby. The agent could move with the user, jumping from hardware to hardware as needed.

A conversation with such a mobile agent could be structured completely different from a fixed agent. It could behave more like an assistant or friend guiding the user instead of a receptionist directing the user. In the end, perhaps we are trying too hard to imitate the receptionist we know instead of finding our own way with all the possibilities our technology offers.

6 Conclusions and Future Work

The experiment used a speech interface for interacting with users. While a lower error rate would be preferable, the overall success rate indicates that the agent can recover from these errors. This shows that technology and voice recognition have evolved far enough to be completely viable even in a noisy environment like a corporate lobby. Break-down in communication was rare and in most cases could be recovered from by repeating the last communication. From the users perspective the experience was neutral, neither positive, nor negative.

The main goal of the experiment was to test to what extent a body contributes to a better user experience with an automated receptionist. A comparison between the two setups indicated no significant difference in the reported user experience between a smart display and a Pepper robot.

The results are inconclusive. This leaves us in uncertainty as to the actual possible benefit of having a body. Given these results we can not recommend investing in expensive Pepper units for commercial automated receptions at scale at this time. The smart displays offer a viable alternative at much less cost and risk. The display can be upgraded to actual robots at a later point in time without serious loss of time and money.

Further research would be needed to reach a clearer picture on the actual benefits of having a body. With the current research design this would take an estimated period of 9 to 12 months assuming the same density of participants.

An other avenue to research would be the effect of agent mobility on the user experience. Combining strategically located smart displays with users smart phones could lead to a different reception experience altogether.

Aside from the mobility of the agent, a locale independent setup would need to change conversation style from the current version. In this experiment it mimics the actual conversation between a human receptionist and a guest. How a conversation could be structured if it no longer followed human conventions and limitations is a question that remains open.

One thing is clear, we have the technology to automate receptions and, at least in regard to our participants, guest are willing to use an automated reception. Implementing an reception agent will alter an organisation. In the extreme case of a fully mobile agent conventional lobbies and receptions desk may disappear as receiving guest becomes an ambient task of the organisation as a whole. The social impact of the could be far reaching and needs to be examined.

Acknowledgments. The authors would like to thank Joost Bosman, Edwin van Dillen and Martijn Schuts for their willing contributions to this project.

References

1. Amazon.com Inc.: Introducing echo show 5: compact smart display with alexa (2019). https://www.amazon.com/gp/product/B07HZLHPKP. Accessed 12 June 2019
2. Baba, A., Kagehiro, T., Koshizuka, H., Togami, M., Yoshiuchi, H.: Robotics solutions opening up new service markets. Hitachi Rev. **65**(9), 433 (2016)
3. Bainbridge, W.A., Hart, J.W., Kim, E.S., Scassellati, B.: The benefits of interactions with physically present robots over video-displayed agents. Int. J. Soc. Robot. **3**(1), 41–52 (2011)
4. Berry, L.L., Wall, E.A., Carbone, L.P.: Service clues and customer assessment of the service experience: lessons from marketing. Acad. Manag. Perspect. **20**(2), 43–57 (2006)
5. Bohus, D., Rudnicky, A.I.: Sorry, I didn't catch that!-an investigation of non-understanding errors and recovery strategies. In: 6th SIGdial Workshop on Discourse and Dialogue (2005)
6. Breazeal, C.: Social interactions in HRI: the robot view. IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) **34**(2), 181–186 (2004)
7. Donahue, T.J., Scheutz, M.: Investigating the effects of robot affect and embodiment on attention and natural language of human teammates. In: 2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 397–402. IEEE (2015)
8. Dunford, R., Palmer, I., Benveniste, J.: Business model replication for early and rapid internationalisation: the ING direct experience. Long Range Plan. **43**(5–6), 655–674 (2010)
9. Dutta, V.: Banking revisited: key trends reshaping banking in India. Paradigm **7**(1), 103–108 (2003)
10. Etlinger, S., Altimeter, A.: The conversational business (2017). Accessed 14 July 2017
11. Hanson, D.: Commencement 2018 keynote address—David Hanson via robotic proxy Sophia (2018)
12. Hashimoto, T., Kobayashi, H.: Study on natural head motion in waiting state with receptionist robot SAYA that has human-like appearance. In: 2009 IEEE Workshop on Robotic Intelligence in Informationally Structured Space, pp. 93–98. IEEE (2009)
13. Julian, C.C., Ramaseshan, B.: The role of customer-contact personnel in the marketing of a retail bank's services. Int. J. Retail Distrib. Manag. **22**(5), 29–34 (1994)
14. Kennedy, J., Baxter, P., Belpaeme, T.: The robot who tried too hard: social behaviour of a robot tutor can negatively affect child learning. In: 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 67–74. IEEE (2015)
15. Lee, K.M., Jung, Y., Kim, J., Kim, S.R.: Are physically embodied social agents better than disembodied social agents?: the effects of physical embodiment, tactile interaction, and people's loneliness in human-robot interaction. Int. J. Hum. Comput. Stud. **64**(10), 962–973 (2006)

16. Lee, M.K., Kiesler, S., Forlizzi, J.: Receptionist or information kiosk: how do people talk with a robot? In: Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, pp. 31–40. ACM (2010)
17. Linssen, J., Theune, M.: R3D3: the rolling receptionist robot with double Dutch dialogue. In: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, pp. 189–190. ACM (2017)
18. Malhotra, R., Malhotra, D.: The impact of internet and e-commerce on the evolving business models in the financial services industry. Int. J. Electron. Bus. **4**(1), 56–82 (2006)
19. Ogawa, K., Nishio, S., Koda, K., Balistreri, G., Watanabe, T., Ishiguro, H.: Exploring the natural reaction of young and aged person with telenoid in a real world. JACIII **15**(5), 592–597 (2011)
20. SoftBank Robotics: Pepper (2016). https://www.softbankrobotics.com/emea/en/pepper. Accessed 12 June 2019
21. Swaid, S.I., Wigand, R.T.: The effect of perceived site-to-store service quality on perceived value and loyalty intentions in multichannel retailing. Int. J. Manag. **29**(3), 301 (2012)
22. Thellman, S., Silvervarg, A., Gulz, A., Ziemke, T.: Physical vs. virtual agent embodiment and effects on social interaction. In: Traum, D., Swartout, W., Khooshabeh, P., Kopp, S., Scherer, S., Leuski, A. (eds.) IVA 2016. LNCS (LNAI), vol. 10011, pp. 412–415. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47665-0_44
23. Tuzovic, S., Paluch, S.: Conversational commerce – a new era for service business development? In: Bruhn, M., Hadwich, K. (eds.) Service Business Development, pp. 81–100. Springer, Wiesbaden (2018). https://doi.org/10.1007/978-3-658-22426-4_4
24. Wainer, J., Feil-Seifer, D.J., Shell, D.A., Mataric, M.J.: The role of physical embodiment in human-robot interaction. In: The 15th IEEE International Symposium on Robot and Human Interactive Communication, ROMAN 2006, pp. 117–122. IEEE (2006)
25. Weiss, A., Bernhaupt, R., Lankes, M., Tscheligi, M.: The USUS evaluation framework for human-robot interaction. In: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction, AISB 2009, vol. 4, pp. 11–26 (2009)
26. Weiss, A., Bernhaupt, R., Tscheligi, M., Yoshida, E.: Addressing user experience and societal impact in a user study with a humanoid robot. In: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction, AISB 2009 (Edinburgh, 8–9 April 2009), SSAISB, pp. 150–157. Citeseer (2009)